

From primitive spaces of bounded rank matrices to a generalized Gerstenhaber theorem

Clément de Seguins Pazzis^{*†}

November 22, 2012

Abstract

A recent generalization of Gerstenhaber's theorem on spaces of nilpotent matrices is derived, under mild conditions on the cardinality of the underlying field, from Atkinson's structure theorem on primitive spaces of bounded rank matrices.

AMS Classification : 15A30; 15A03

Keywords : rank, nilpotency, spectrum, transitivity, vector spaces of matrices

1 Introduction

In the modern geometric theory of matrices, two deep structure theorems stand out: Dieudonné's theorem [4] on large spaces of singular matrices - later generalized by Atkinson and Lloyd [3] - and Gerstenhaber's theorem on spaces of nilpotent matrices [5]. Very recent advances have been made in both problems: Atkinson and Lloyd's extension of Dieudonné's theorem has been shown to hold for almost all fields [10], while there have been several generalizations of Gerstenhaber's theorem, most notably to **trivial spectrum** spaces of matrices, i.e., subspaces of square matrices having no non-zero eigenvalue in their field of definition.

^{*}Université de Versailles Saint-Quentin-en-Yvelines, Laboratoire de Mathématiques de Versailles, 45 avenue des Etats-Unis, 78035 Versailles cedex, France

[†]e-mail address: dsp.prof@gmail.com

In our view, another structure theorem has not received the attention it deserved from the mathematics community: it is Atkinson's theorem on primitive spaces of bounded rank matrices [2, Theorem B]. In this note, we will expose a yet unknown connection between Atkinson's theorem and Gerstenhaber's, and will use this insight to obtain a greatly simplified proof of a slightly weaker version of the classification theorem for trivial spectrum spaces of matrices. We believe that the technique displayed here has a great potential to deliver new insights into the structure of spaces of nilpotent matrices, as it only uses a limit case in Atkinson's theorem.

2 Notation and basic definition

Here, \mathbb{K} denotes an arbitrary field, and $M_{m,n}(\mathbb{K})$, $M_n(\mathbb{K})$, $A_n(\mathbb{K})$, $GL_n(\mathbb{K})$ respectively denote the sets of all $m \times n$ matrices, $n \times n$ matrices, $n \times n$ alternating matrices, and $n \times n$ invertible matrices with entries in \mathbb{K} . Two subsets \mathcal{V} and \mathcal{W} of $M_{m,n}(\mathbb{K})$ are called equivalent when there exists a pair $(P, Q) \in GL_m(\mathbb{K}) \times GL_n(\mathbb{K})$ such that $\mathcal{V} = P\mathcal{W}Q$, which means that \mathcal{V} and \mathcal{W} represent the same set of linear operators in different choices of bases of the source and goal spaces. If $m = n$, we say that \mathcal{V} and \mathcal{W} are similar when, in the above condition, we require that $Q = P^{-1}$, meaning that \mathcal{V} and \mathcal{W} represent the same set of endomorphisms of a vector space in two potentially different bases.

Given a subset \mathcal{V} of $M_{m,n}(\mathbb{K})$, the upper rank of \mathcal{V} , denoted by $urk(\mathcal{V})$, is the largest rank for a matrix in \mathcal{V} . Note that two equivalent subsets share the same upper rank.

A linear subspace \mathcal{V} of $M_{m,n}(\mathbb{K})$ is called **primitive** when it satisfies the following four conditions:

- (i) \mathcal{V} is not equivalent to a space of matrices with the last column equal to zero;
- (ii) \mathcal{V} is not equivalent to a space of matrices with the last row equal to zero;
- (iii) \mathcal{V} is not equivalent to a space \mathcal{V}' in which every matrix is written as $M = [H(M) \quad ?]_{m \times 1}$ and $urk H(\mathcal{V}') < urk \mathcal{V}$;
- (iv) \mathcal{V} is not equivalent to a space \mathcal{V}' in which every matrix is written as $M = \begin{bmatrix} R(M) \\ ? \end{bmatrix}_{1 \times n}$ and $urk R(\mathcal{V}') < urk \mathcal{V}$.

Moreover, we say that \mathcal{V} is **semi-primitive** when it is only required to satisfy conditions (i), (ii) and (iii); we say that \mathcal{V} is **reduced** when it is only required to satisfy conditions (i) and (ii).

We note that the upper rank of a semi-primitive subspace of $M_{m,n}(\mathbb{K})$ is always less than n , so that the upper rank of a primitive subspace is always less than m and n . As shown by the first statement in Theorem 1 of [1], the primitive spaces are the elementary pieces upon which are built all those matrix spaces with upper rank less than the number of rows and the number of columns. A fundamental example of primitive space can be derived from the canonical pairing

$$\varphi_n : \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}^n \wedge \mathbb{K}^n.$$

In the canonical basis (e_1, \dots, e_n) of \mathbb{K}^n and the lexicographically ordered basis $(e_i \wedge e_j)_{1 \leq i < j \leq n}$ of $\mathbb{K}^n \wedge \mathbb{K}^n$, one then takes the space $S(\varphi_n)$ of all matrices representing the operators $x \wedge -$ for $x \in \mathbb{K}^n$. One checks that, for $n \geq 2$, $S(\varphi_n)$ is a semi-primitive linear subspace of $M_{\binom{n}{2}, n}(\mathbb{K})$ - and even a primitive one if $n > 2$ - and that it is also the space of all matrices

$$\begin{bmatrix} X^T A_1 \\ \vdots \\ X^T A_{\binom{n}{2}} \end{bmatrix} \quad \text{with } X \in \mathbb{K}^n,$$

where the matrices $A_1, \dots, A_{\binom{n}{2}}$ are the elements of $(e_i e_j^T - e_j e_i^T)_{1 \leq i < j \leq n}$ put in the lexicographical order. A subspace \mathcal{V} equivalent to $S(\varphi_n)$ is exactly a subspace for which there is a basis $(B_1, \dots, B_{\binom{n}{2}})$ of $A_n(\mathbb{K})$ together with some $P \in GL_n(\mathbb{K})$ such that \mathcal{V} is the space of all matrices

$$\begin{bmatrix} X^T B_1 P \\ \vdots \\ X^T B_{\binom{n}{2}} P \end{bmatrix} \quad \text{with } X \in \mathbb{K}^n.$$

3 Atkinson's theorem

We now state a special case of the transposed version of Atkinson's theorem, combining Lemma 6 of [1] and Theorem B of [2].

Theorem 1 (Atkinson). *Let \mathcal{S} be a semi-primitive linear subspace of $M_{m,n}(\mathbb{K})$. Set $r := \text{urk } \mathcal{S}$ and assume that $\#\mathbb{K} > r$. Then, $m \leq \frac{r(r+1)}{2}$. If in addition $m = \frac{r(r+1)}{2}$ and $r > 1$, then $n = r + 1$ and \mathcal{S} is equivalent to $S(\varphi_n)$.*

Note that, for the case $m = \frac{r(r+1)}{2}$ and $r > 1$, Atkinson only states his result for primitive spaces: however, if we put Atkinson's arguments in the context of our version (i.e., we transpose them), then the only instance when he uses condition (iv) is to discard the case where \mathcal{S} might be equivalent to a space \mathcal{S}' of matrices of the form $M = \begin{bmatrix} [?]_{1 \times q} & [?]_{1 \times (n-q)} \\ H(M) & [0]_{(m-1) \times (n-q)} \end{bmatrix}$ with $q \in \llbracket 0, n-1 \rrbracket$ and a reduced space $H(\mathcal{S}')$. But then $\text{urk } H(\mathcal{S}') \leq r-1$ and Lemma 6 of [1] would show that $m-1 \leq \binom{r}{2} < \binom{r+1}{2} - 1$ as, in the terminology of Atkinson and Lloyd, the space $H(\mathcal{S}')$ would have the column property.

We also remark that the conclusion of the second statement still holds in the case $m = 1$ and $n = 2$, as then $\mathcal{S} = M_{1,2}(\mathbb{K}) = S(\varphi_2)$.

4 The generalized Gerstenhaber theorem

Seemingly unrelated to Atkinson's theory of primitive spaces is the following generalization of Gerstenhaber's theorem, proved in [8], in which a linear subspace \mathcal{V} of $M_n(\mathbb{K})$ is called **irreducible** if no proper non-zero linear subspace of \mathbb{K}^n is globally invariant under all the elements of \mathcal{V} .

Theorem 2 (Generalized Gerstenhaber theorem). *Let \mathcal{V} be a trivial spectrum linear subspace of $M_n(\mathbb{K})$. Then,*

$$\dim \mathcal{V} \leq \frac{n(n-1)}{2}. \quad (1)$$

If equality holds in (1) with $\#\mathbb{K} \geq 3$ and \mathcal{V} irreducible, then $\mathcal{V} = P A_n(\mathbb{K})$ for some $P \in \text{GL}_n(\mathbb{K})$.

Given $P \in \text{GL}_n(\mathbb{K})$, one then shows that $P A_n(\mathbb{K})$ is an irreducible space with a trivial spectrum if and only if $X \mapsto X^T P X$ is a non-isotropic quadratic

form on \mathbb{K}^n , see [8, Lemma 10]. Theorem 2 may be used to provide a complete classification of trivial spectrum spaces with the maximal dimension - up to similarity - reducing the problem to the classification of non-isotropic bilinear forms [8, Theorem 4] up to similarity.

The main rationale for studying trivial spectrum spaces stems from their ties to affine spaces of non-singular matrices. In short, if \mathcal{A} is an affine subspace of non-singular matrices of $M_n(\mathbb{K})$ which contains I_n , then its translation vector space $\mathcal{A} - I_n$ is a trivial spectrum linear subspace. With little effort, Theorem 2 shows that any such affine subspace has dimension less than or equal to $\binom{n}{2}$, and reduces the classification - up to equivalence - of those with the maximal dimension to the classification of non-isotropic quadratic forms on \mathbb{K} , up to similarity [8, Theorem 7]. This has led to similar structure theorems for large affine spaces of matrices with a lower bound on the rank [7].

Finally, Theorem 2 yields Gerstenhaber's theorem - which we recall now - as an easy consequence for fields with more than 2 elements (for proofs that hold regardless of the cardinality of the field, see [9] and [11], while an elegant proof that works only for fields with more than two elements is featured in [6]).

Theorem 3 (Gerstenhaber's theorem [5]). *Let \mathcal{V} be a linear subspace of $M_n(\mathbb{K})$ in which all the matrices are nilpotent. Then, $\dim \mathcal{V} \leq \frac{n(n-1)}{2}$, and equality holds if and only if \mathcal{V} is similar to the space of all strictly upper-triangular $n \times n$ matrices.*

5 Proof of the generalized Gerstenhaber theorem in a restricted setting

Spread over twenty dense pages, the only known proof of Theorem 2 is a very intricate tour de force. Using Atkinson's theorem on semi-primitive spaces, we shall now give a much shorter alternative proof *under the additional assumption that $\#\mathbb{K} \geq n$* .

Let us immediately explain the connection with Atkinson's theorem. Let \mathcal{V} be a trivial spectrum subspace of $M_n(\mathbb{K})$. Then, \mathcal{V} has an interesting property that was the cornerstone of the proof of Theorem 2 featured in [8]: it is **totally intransitive** in the sense that, for every non-zero vector $X \in \mathbb{K}^n$, the linear subspace $\mathcal{V}X := \{NX \mid N \in \mathcal{V}\}$ is a proper subspace of \mathbb{K}^n as it cannot contain X .

We now combine this simple fact with a duality argument to create a linear

subspace of $M_{m,n}(\mathbb{K})$ with upper rank less than n , where $m := \dim \mathcal{V}$. For $X \in \mathbb{K}^n$, consider the bilinear form

$$\widehat{X} : (N, Y) \in \mathcal{V} \times \mathbb{K}^n \mapsto Y^T N X \in \mathbb{K},$$

and denote by $\mathcal{M} \subset M_{m,n}(\mathbb{K})$ the space of all matrices representing such forms in chosen bases of \mathcal{V} and \mathbb{K}^n . In other words, if the chosen basis of \mathcal{V} is (B_1, \dots, B_m) , then there is an invertible matrix $P \in GL_n(\mathbb{K})$ such that, for all $X \in \mathbb{K}^n$, the matrix of \widehat{X} is

$$\begin{bmatrix} X^T B_1^T P \\ \vdots \\ X^T B_m^T P \end{bmatrix}.$$

Note that \mathcal{V} always satisfies condition (ii), as a matrix $N \in M_n(\mathbb{K})$ which satisfies $Y^T N X = 0$ for all $(X, Y) \in (\mathbb{K}^n)^2$ is necessarily zero. We also see that $\text{urk } \mathcal{V} < n$ since, given a non-zero vector $X \in \mathbb{K}^n$, the set $\mathcal{V}X$ is a proper linear subspace of \mathbb{K}^n , which yields a non-zero vector $Y \in \mathbb{K}^n$ for which $Y^T \mathcal{V}X = 0$.

Now, it may very well happen that \mathcal{M} is not semi-primitive, but let us assume for the moment that it is. Then, by Atkinson's theorem, we would have $m \leq \binom{n}{2}$, and, in case of equality, we would find a basis (A_1, \dots, A_m) of $A_n(\mathbb{K})$ together with an invertible matrix $Q \in GL_n(\mathbb{K})$ such that

$$\forall X \in \mathbb{K}^n, \begin{bmatrix} X^T B_1^T P \\ \vdots \\ X^T B_m^T P \end{bmatrix} = \begin{bmatrix} X^T A_1 Q \\ \vdots \\ X^T A_m Q \end{bmatrix};$$

this would yield $B_i^T P = A_i Q$ for all $i \in [\![1, m]\!]$, and hence we would conclude that

$$\mathcal{V} = (QP^{-1})^T A_n(\mathbb{K}).$$

Let us now return to the general case. The proof works by induction on n . The result is trivial for $n = 1$ and we now assume that $n \geq 2$. Note first that we may always assume that \mathcal{V} is irreducible, for if it is not, then, replacing \mathcal{V} by a similar subspace, we can see that no generality is lost in assuming that there exists an integer $p \in [\![1, n - 1]\!]$ such that every matrix of \mathcal{V} splits up as

$$N = \begin{bmatrix} A(N) & [?]_{p \times (n-p)} \\ [0] & B(N) \end{bmatrix}$$

where $A(N)$ and $B(N)$ are respectively $p \times p$ and $(n-p) \times (n-p)$ matrices; $A(\mathcal{V})$ and $B(\mathcal{V})$ are then trivial spectrum subspaces respectively of $M_p(\mathbb{K})$ and $M_{n-p}(\mathbb{K})$, and by induction we deduce that

$$\dim \mathcal{V} \leq \binom{p}{2} + \binom{n-p}{2} + p(n-p) = \binom{n}{2}.$$

In the rest of the proof, we assume that \mathcal{V} is irreducible.

Let us come back to the matrix space \mathcal{M} . It now satisfies condition (i), for if it did not, then we would have a non-zero vector $Y \in \mathbb{K}^n$ for which $\forall X \in \mathbb{K}^n, \forall N \in \mathcal{V}, Y^T N X = 0$, and hence all the elements of \mathcal{V} would map all the vectors of \mathbb{K}^n into some fixed linear hyperplane of \mathbb{K}^n , contradicting the assumed irreducibility of \mathcal{V} . Therefore, \mathcal{M} satisfies conditions (i), (ii), and $\text{urk } \mathcal{M} < n$. If \mathcal{M} is semi-primitive, then the conclusion follows as explained earlier.

If we now assume that \mathcal{M} is not semi-primitive, then we may then find a *minimal* integer $d \in \llbracket 1, n-1 \rrbracket$ for which the bases of \mathcal{V} and \mathbb{K}^n are chosen so that every matrix of \mathcal{M} has the form

$$M = [H(M) \quad ?]_{m \times (n-d)},$$

and $H(\mathcal{M})$ is a linear subspace of $M_{m,d}(\mathbb{K})$ with upper rank less than d . In this situation, we may also modify the basis of \mathcal{V} further to the point where, for every $M \in \mathcal{M}$, we have

$$H(M) = \begin{bmatrix} K(M) \\ [0]_{(m-c) \times d} \end{bmatrix},$$

where $K(\mathcal{M})$ is a linear subspace of $M_{c,d}(\mathbb{K})$ with upper rank less than d and which satisfies condition (ii). Then, using the minimality of d and the fact that \mathcal{M} satisfies condition (i), we can see that $K(\mathcal{M})$ is a semi-primitive subspace of $M_{c,d}(\mathbb{K})$. Theorem 1 then yields

$$c \leq \binom{d}{2}.$$

Let us now come back to \mathcal{V} and see how the above reduction plays out. As we may replace \mathcal{V} with a similar subspace, no generality is lost in assuming that

the chosen basis of \mathbb{K}^n is the canonical one. Let us then write every matrix N of \mathcal{V} as

$$N = \begin{bmatrix} R_1(N) \\ R_2(N) \end{bmatrix} = \begin{bmatrix} A(N) & C(N) \\ B(N) & D(N) \end{bmatrix},$$

where $R_1(N)$, $R_2(N)$, $A(N)$, $B(N)$, $C(N)$ and $D(N)$ are respectively $d \times n$, $(n-d) \times n$, $d \times d$, $(n-d) \times d$, $d \times (n-d)$ and $(n-d) \times (n-d)$ matrices. Set also

$$\mathcal{W} := \text{Ker } R_1.$$

As $K(\mathcal{M})$ satisfies condition (ii), one sees that the last $m - c$ vectors of the chosen basis of \mathcal{V} span the subspace of all matrices of \mathcal{V} in which the first d rows equal zero, and hence

$$m - c = \dim \mathcal{W}.$$

On the other hand, every matrix N of \mathcal{W} splits up as

$$N = \begin{bmatrix} [0] \\ B(N) \end{bmatrix} \quad \begin{bmatrix} [0] \\ D(N) \end{bmatrix},$$

and hence $D(\mathcal{W})$ is a trivial spectrum subspace of $M_{n-d}(\mathbb{K})$. By induction, we have $\dim D(\mathcal{W}) \leq \binom{n-d}{2}$, and therefore

$$\dim \mathcal{W} \leq (n-d)d + \binom{n-d}{2}.$$

We conclude that

$$m = c + \dim \mathcal{W} \leq \binom{d}{2} + (n-d)d + \binom{n-d}{2} = \binom{n}{2},$$

thus completing the proof of inequality (1).

If we now assume that $m = \binom{n}{2}$ on top of the previous assumptions, then all the above inequalities turn out to be equalities, and in particular we have:

$$(a) \quad \dim D(\mathcal{W}) = \binom{n-d}{2};$$

(b) The space \mathcal{V} contains every matrix of the form

$$\begin{bmatrix} [0]_{d \times d} & [0]_{d \times (n-d)} \\ [?]_{(n-d) \times d} & [0]_{(n-d) \times (n-d)} \end{bmatrix}.$$

Using point (b), we deduce that, for every $N \in \mathcal{V}$, the matrix

$$\begin{bmatrix} A(N) & C(N) \\ [0] & D(N) \end{bmatrix}$$

belongs to \mathcal{V} , and hence $D(\mathcal{V})$ is a trivial spectrum subspace of $M_{n-d}(\mathbb{K})$. By induction, we have $\dim D(\mathcal{V}) \leq \binom{n-d}{2} = \dim D(\mathcal{W})$ with $D(\mathcal{W}) \subset D(\mathcal{V})$, which yields $D(\mathcal{W}) = D(\mathcal{V})$.

We shall now obtain a contradiction from an invariance argument. Assume that some $N_0 \in \mathcal{V}$ satisfies $C(N_0) \neq 0$. Then, we may find a non-zero vector $x \in \mathbb{K}^{n-d}$ such that $C(N_0)x \neq 0$, and then choose $R \in M_{n-d,d}(\mathbb{K})$ for which

$RC(N_0)x = x$. With the invertible matrix $P := \begin{bmatrix} I_d & [0] \\ R & I_{n-d} \end{bmatrix}$, one computes that

$$\forall N \in \mathcal{V}, \quad PNP^{-1} = \begin{bmatrix} A(N) - C(N)R & C(N) \\ B(N) + RA(N) - RC(N)R & D(N) + RC(N) \end{bmatrix}.$$

In the new trivial spectrum space $\mathcal{V}' := P\mathcal{V}P^{-1}$, the subspace of matrices with all first d rows equal to zero is still \mathcal{W} . Thus, with the above dimensional arguments applied to \mathcal{V}' , we can deduce that $D(N_0) + RC(N_0)$ belongs to $D(\mathcal{V}') = D(\mathcal{W}) = D(\mathcal{V})$, and hence $RC(N_0)$ belongs to $D(\mathcal{V})$. This is absurd because we have seen that $D(\mathcal{V})$ is a trivial spectrum subspace of $M_{n-d}(\mathbb{K})$.

Therefore, $C(N) = 0$ for all $N \in \mathcal{V}$, and hence $\{0\} \times \mathbb{K}^{n-d}$ is a globally invariant subspace for all the matrices of \mathcal{V} . This contradicts the assumed irreducibility of \mathcal{V} and concludes the proof.

References

- [1] M.D. Atkinson, S. Lloyd, Primitive spaces of matrices of bounded rank, *J. Austral. Math. Soc. (Ser. A)* **30** (1980) 473-482.
- [2] M.D. Atkinson, Primitive spaces of matrices of bounded rank II, *J. Austral. Math. Soc. (Ser. A)* **34** (1983) 306-315.
- [3] M.D. Atkinson, S. Lloyd, Large spaces of matrices of bounded rank, *Quart. J. Math. Oxford (2)* **31** (1980) 253-262.
- [4] J. Dieudonné, Sur une généralisation du groupe orthogonal à quatre variables, *Arch. Math.* **1** (1949) 282-287.

- [5] M. Gerstenhaber, On Nilalgebras and Linear Varieties of Nilpotent Matrices (I), *Amer. J. Math.* **80** (1958) 614-622.
- [6] B. Mathes, M. Omladič, H. Radjavi, Linear spaces of nilpotent matrices, *Linear Algebra Appl.* **149** (1991) 215-225.
- [7] C. de Seguins Pazzis, Large affine spaces of matrices with rank bounded below, *Linear Algebra Appl.* **437-2** (2012) 499-518.
- [8] C. de Seguins Pazzis, Large affine spaces of non-singular matrices, *Trans. Amer. Math. Soc.*, in press, *arXiv preprint* <http://arxiv.org/abs/1102.2493>
- [9] C. de Seguins Pazzis, On Gerstenhaber's theorem for spaces of nilpotent matrices over a skew field, *arXiv preprint* <http://arxiv.org/abs/1210.4951>
- [10] C. de Seguins Pazzis, The classification of large spaces of matrices with bounded rank, *arXiv preprint* <http://arxiv.org/abs/1004.0298>
- [11] V.N. Serezhkin, Linear transformations preserving nilpotency (in Russian), *Izv. Akad. Nauk BSSR, Ser. Fiz.-Mat. Nauk* **125** (1985) 46-50.